

SEQUENCE LISTING

<110> Pharmexa A/S
<120> PURIFICATION OF HER-2 VARIANTS
<130> P1020DK00
<160> 2
<170> PatentIn version 3.2
<210> 1
<211> 5661
<212> DNA
<213> Artificial sequence
<220>
<223> Recombinant expression plasmid derived from pMT

<220>
<221> polyA_signal
<222> (263)..(268)
<223> SV40 late polyadenylation site

<220>
<221> misc_feature
<222> (1579)..(2439)
<223> Ampicillin resistance gene, encoded by complementary strand

<220>
<221> promoter
<222> (3050)..(3415)
<223> Metallothionein promoter

<220>
<221> RBS
<222> (3493)..(3501)
<223> Kozak-like sequence

<220>
<221> CDS
<222> (3502)..(5592)
<223> DNA encoding immunogenic, his-tagged variant of human HER-2

<220>
<221> sig_peptide
<222> (3502)..(3555)
<223> BiP signal sequence

<220>
<221> misc_feature
<222> (3556)..(3597)
<223> Histidine tag

<220>
<221> mat_peptide
<222> (3556)..(5589)

```

<220>
<221> misc_feature
<222> (3598)..(3603)
<223> Dipeptidase stop sequence

<220>
<221> misc_feature
<222> (3604)..(5589)
<223> Gene coding for the hHER2MA5-5DUH protein

<220>
<221> misc_feature
<222> (4357)..(4401)
<223> Diphtheria toxoid P2 epitope

<220>
<221> misc_feature
<222> (5500)..(5562)
<223> Diphtheria toxoid P30 epitope

<400> 1
ggccgctcga gtctagaggg cccttcgaag gtaaggctat ccctaaccct ctcctcggtc      60
tcgattctac gcgtaccggt catcatcacc atcaccattg agtttaaacc cgctgatcag
cctcgactgt gccttctaag atccagacat gataagatac attgatgagt ttggacaaac      120
cacaactaga atgcagtgaa aaaaatgctt tatttgcgaa atttgcgtg ctattgcctt      180
atttgcgtt accataagct gcaataaaaca agttaacaac aacaattgca ttcattttat      240
gtttcagggtt cagggggagg tgtgggaggt tttttaaagc aagtaaaacc tctacaaatg      300
tggtatggct gattatgatc agtcgacctg caggcatgca agcttggcgt aatcatggtc      360
atagctgtt cctgtgtgaa attgttatcc gctcacaatt ccacacaaca tacgagccgg      420
aagcataaaag tgtaaagcct ggggtgccta atgagtgagc taactcacat taattgcgtt      480
gcgcgtcaactg cccgcattcc agtcggaaaa cctgtcggtc cagctgcatt aatgaatcgg      540
ccaaacgcgcg gggagaggcg gtttgcgtat tggcgctct tccgcttcct cgctcaactga      600
ctcgctgcgc tcggctcggtc ggctgcggcg agcggtatca gctcactcaa aggcggtaat      660
acggttatcc acagaatcag gggataacgc aggaaagaac atgtgagcaa aaggccagca      720
aaaggccagg aaccgtaaaa aggccgcgtt gctggcggtt ttccataggc tccggccccc      780
tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga caggactata      840
aagataccag gcgttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc      900
gcttaccgga tacctgtccg ccttctccc ttccggaaagc gtggcgcttt ctcatagctc      960
acgctgtagg tatctcagtt cggtgttaggt cgttcgctcc aagctgggct gtgtgcacga     1020
accccccgtt cagcccgacc gctgcgcctt atcccgtaac tatcgtcttg agtccaaccc     1080
ggtaagacac gacttatcgc cactggcagc agccactggc aacaggatta gcagagcag     1140
ggtaagacac gacttatcgc cactggcagc agccactggc aacaggatta gcagagcag     1200

```

gtatgttaggc ggtgctacag agttcttcaa gtgggtggcct aactacggct acactagaag	1260
gacagtattt ggtatctgcg ctctgctgaa gccagttacc ttccggaaaaa gagttggtag	1320
ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt tttttgttt gcaaggcagca	1380
gattacgcgc agaaaaaaag gatctcaaga agatcctttg atctttctta cggggctctga	1440
cgctcagtgg aacgaaaact cacgttaagg gatTTTggtc atgagattat caaaaaggat	1500
cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga	1560
gtaaaacttgg tctgacagtt accaatgctt aatcagttag gcacccatct cagcgatctg	1620
tctatttcgt tcatccatag ttgcctgact ccccgtcgtg tagataacta cgatacggga	1680
gggcttacca tctggccccca gtgctgcaat gataccgcga gacccacgct caccggctcc	1740
agatttatca gcaataaacc agccagccgg aaggggccgag cgcagaagtg gtcctgcaac	1800
tttacccgcc tccatccagt ctattaattt tgccggaa gctagagtaa gtagttcgcc	1860
agttaatagt ttgcgcacg ttgttgcctt tgctacaggc atcgtggtgt caccgctcgtc	1920
gtttggatg gtttcattca gtcgggttc ccaacgatca aggcgagttt catgatcccc	1980
catgttgc aaaaaagcgg ttagctcctt cggccctccg atcgttgtca gaagtaagtt	2040
ggccgcagtg ttatcactca tggttatggc agcaactgcat aattctctta ctgtcatgcc	2100
atccgttaaga tgctttctg tgactggta gtaactcaacc aagtcatctt gagaatagt	2160
tatgcggcga ccgagttgtc ttgcggcgtc gtcaatacgg gataataccg cgccacatag	2220
cagaacttta aaagtgcac tcattggaaa acgttcttcg gggcgaaaaac tctcaaggat	2280
cttaccgctg ttgagatcca gttcgatgtc acccaactcgt gcacccaaact gatcttcagc	2340
atctttact ttcaccagcg tttctgggtg agcaaaaaaca ggaaggcaaa atgcccggaaa	2400
aaagggaata agggcgacac ggaaatgtt aataactcata ctcttccttt ttcaatattt	2460
ttgaagcatt tattcagggtt attgtctcat gagcggatac atatggat gtatggaa	2520
aaataaaacaa ataggggttc cgccacatt tccccggaaa gtgccacctg acgtctaaga	2580
aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc cctttcggttc	2640
gcgcgtttcg gtgatgacgg tgaaaacctc tgacacatgc agctccggaa gacggtcaca	2700
gcttgcgtgt aagcgatgc cgggagcaga caagccgtc agggcggtc agcgggtgtt	2760
ggcggtgtc gggcgatggct taactatgcg gcatcagagc agattgtact gagagtgcac	2820
catatgcgtt gtgaaatacc gcacagatgc gtaaggagaa aataccgcat caggcgccat	2880
tcgcccattca ggctgcgcaa ctgttggaa gggcgatcgg tgccggcctc ttgcgttattt	2940
cgccagctgg cgaaaggggg atgtgctgca aggcgattaa gttggtaac gccagggttt	3000

tcccagtcac gacgttgtaa aacgacggcc agtgcgcagt aattaattcg ttgcaggaca. 3060
 ggatgtggtg cccgatgtga cttagctttt gctgcaggcc gtcctatcct ctggttccga 3120
 taagagaccc agaactccgg ccccccacccg cccaccgcca cccccataca tatgtggtac 3180
 gcaagtaaga gtgcctgcgc atgccccatg tgccccacca agagtttgc atcccataca 3240
 agtccccaaa gtggagaacc gaaccaattc ttgcgggca gaacaaaagc ttctgcacac 3300
 gtctccactc gaatttggag ccggccggcg ttgtcaaaag aggtgaatcg aacgaaaagac 3360
 ccgtgtgtaa agccgcgttt ccaaaatgtta taaaaccgag agcatctggc caatgtgcat 3420
 cagttgtgtt cagcagcaaa atcaagtgaa tcatctcagt gcaactaaag gggggatcta 3480
 gatcggggta ccaaagtcac c atg aag ttg tgc atc ttg ctg gcc gtc gtg 3531
 Met Lys Leu Cys Ile Leu Leu Ala Val Val
 -15 -10
 gcc ttc gtg ggc ctg tcg ctg ggc atg aag cac caa cac caa cat caa 3579
 Ala Phe Val Gly Leu Ser Leu Gly Met Lys His Gln His Gln His Gln
 -5 -1 1 5
 cat caa cat caa cat caa gcc ccc tcc acc caa gtg tgt acc ggc aca 3627
 His Gln His Gln His Gln Ala Pro Ser Thr Gln Val Cys Thr Gly Thr
 10 15 20
 gac atg aag ctg cgg ctc cct gcc agt ccc gag acc cac ctg gac atg 3675
 Asp Met Lys Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met
 25 30 35 40
 ctc cgc cac ctc tac cag ggc tgc cag gtg gtg cag gga aac ctg gaa 3723
 Leu Arg His Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu
 45 50 55
 ctc acc tac ctg ccc acc aat gcc agc tta agt ttc ctg cag gat atc 3771
 Leu Thr Tyr Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile
 60 65 70
 cag gag gtg cag ggc tac gtg ctc atc gct cac aac caa gtg agg cag 3819
 Gln Glu Val Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln
 75 80 85
 gtc cca ctg cag agg ctg cgg att gtg cga ggc acc cag ctc ttt gag 3867
 Val Pro Leu Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu
 90 95 100
 gac aac tat gcc ctg gcc gtg cta gac aat gga gac ccc ctg aac aat 3915
 Asp Asn Tyr Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn
 105 110 115 120
 acc acc cct gtc aca ggg gcc tcc cca gga ggc ctg cgg gag ctg cag 3963
 Thr Thr Pro Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln
 125 130 135
 ctt cga agc ctc aca gag atc ttg aaa gga ggg gtc ttg atc cag cgg 4011
 Leu Arg Ser Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg
 140 145 150

aac ccc cag ctc tgc tac cag gac acg att ttg tgg aag gac atc ttc Asn Pro Gln Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe	4059
155 160 165	
cac aag aac aac cag ctg gct ctc aca ctg ata gac acc aac cgc tct His Lys Asn Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser	4107
170 175 180	
cgg gcc tgc cac ccc tgt tct ccg atg tgt aag ggc tcc cgc tgc tgg Arg Ala Cys His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp	4155
185 190 195 200	
gga gag agt tct gag gat tgt cag agc ctg acg cgc act gtc tgt gcc Gly Glu Ser Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala	4203
205 210 215	
ggc ggc tgt gcc cgc tgc aag ggg cca ctg ccc act gac tgc tgc cat Gly Gly Cys Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His	4251
220 225 230	
gag cag tgt gct gcc ggc tgc acg ggc ccc aag cac tct gac tgc ctg Glu Gln Cys Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu	4299
235 240 245	
gcc tgc ctc cac ttc aac cac agt ggc atc tgt gag ctg cac tgc cca Ala Cys Leu His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro	4347
250 255 260	
gcc ctg gtc cag tac atc aaa gct aac tcc aaa ttc atc ggt atc acc Ala Leu Val Gln Tyr Ile Lys Ala Asn Ser Lys Phe Ile Gly Ile Thr	4395
265 270 275 280	
gag ctg cgg tat aca ttc ggc gcc agc tgt gtg act gcc tgt ccc tac Glu Leu Arg Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr	4443
285 290 295	
aac tac ctt tct acg gac gtg gga tcc tgc acc ctc gtc tgc ccc ctg Asn Tyr Leu Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu	4491
300 305 310	
cac aac caa gag gtg aca gca gag gat gga aca cag cgg tgt gag aag His Asn Gln Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys	4539
315 320 325	
tgc agc aag ccc tgt gcc cga gtg tgc tat ggt ctg ggc atg gag cac Cys Ser Lys Pro Cys Ala Arg Val Cys Tyr Gly Leu Gly Met Glu His	4587
330 335 340	
ttg cga gag gtg agg gca gtt acc agt gcc aat atc cag gag ttt gct Leu Arg Glu Val Arg Ala Val Thr Ser Ala Asn Ile Gln Glu Phe Ala	4635
345 350 355 360	
ggc tgc aag aag atc ttt ggg agc ctg gca ttt ctg ccg gag agc ttt Gly Cys Lys Lys Ile Phe Gly Ser Leu Ala Phe Leu Pro Glu Ser Phe	4683
365 370 375	
gat ggg gac cca gcc tcc aac act gcc ccg ctc cag cca gag gag ctc Asp Gly Asp Pro Ala Ser Asn Thr Ala Pro Leu Gln Pro Glu Gln Leu	4731
380 385 390	

caa gtg ttt gag act ctg gaa gag atc aca ggt tac cta tac atc tca Gln Val Phe Glu Thr Leu Glu Glu Ile Thr Gly Tyr Leu Tyr Ile Ser 395 400 405	4779
gca tgg ccg gac agc ctg cct gac ctc agc gtc ttc cag aac ctg caa Ala Trp Pro Asp Ser Leu Pro Asp Leu Ser Val Phe Gln Asn Leu Gln 410 415 420	4827
gta atc cgg gga cga att ctg cac aat ggc gcc tac tcg ctg acc ctg Val Ile Arg Gly Arg Ile Leu His Asn Gly Ala Tyr Ser Leu Thr Leu 425 430 435 440	4875
caa ggg ctg ggc atc agc tgg ctg ggg ctg cgc tca ctg agg gaa ctg Gln Gly Leu Gly Ile Ser Trp Leu Gly Leu Arg Ser Leu Arg Glu Leu 445 450 455	4923
ggc agt gga ctg gcc ctc atc cac cat aac acc cac ctc tgc ttc gtg Gly Ser Gly Leu Ala Leu Ile His His Asn Thr His Leu Cys Phe Val 460 465 470	4971
cac acg gtg ccc tgg gac cag ctc ttt cgg aac ccg cac caa gct ctg His Thr Val Pro Trp Asp Gln Leu Phe Arg Asn Pro His Gln Ala Leu 475 480 485	5019
ctc cac act gcc aac cgg cca gag gac gag tgt gtg ggc gag ggc ctg Leu His Thr Ala Asn Arg Pro Glu Asp Glu Cys Val Gly Glu Gly Leu 490 495 500	5067
gcc tgc cac cag ctg tgc gcc cga ggg cac tgc tgg ggt cca ggg ccc Ala Cys His Gln Leu Cys Ala Arg Gly His Cys Trp Gly Pro Gly Pro 505 510 515 520	5115
acc cag tgt gtc aac tgc agc cag ttc ctt cgg ggc cag gag tgc gtg Thr Gln Cys Val Asn Cys Ser Gln Phe Leu Arg Gly Gln Glu Cys Val 525 530 535	5163
gag gaa tgc cga gta ctg cag ggg ctc ccc agg gag tat gtg aat gcc Glu Glu Cys Arg Val Leu Gln Gly Leu Pro Arg Glu Tyr Val Asn Ala 540 545 550	5211
agg cac tgt ttg ccg tgc cac cct gag tgt cag ccc cag aat ggc tca Arg His Cys Leu Pro Cys His Pro Glu Cys Gln Pro Gln Asn Gly Ser 555 560 565	5259
gtg acc tgt ttt gga ccg gag gct gac cag tgt gtg gcc tgt gcc cac Val Thr Cys Phe Gly Pro Glu Ala Asp Gln Cys Val Ala Cys Ala His 570 575 580	5307
tat aag gac cct ccc ttc tgc gtg gcc cgc tgc ccc agc ggt gtg aaa Tyr Lys Asp Pro Pro Phe Cys Val Ala Arg Cys Pro Ser Gly Val Lys 585 590 595 600	5355
cct gac ctc tcc tac atg ccc atc tgg aag ttt cca gat gag gag ggc Pro Asp Leu Ser Tyr Met Pro Ile Trp Lys Phe Pro Asp Glu Glu Gly 605 610 615	5403
gca tgc cag cct tgc ccc atc aac tgc acc cac tcc tgt gtg gac ctg Ala Cys Gln Pro Cys Pro Ile Asn Cys Thr His Ser Cys Val Asp Leu 620 625 630	5451

gat gac aag ggc tgc ccc gcc gag cag aga gcc agc cct ctg acg tcc 5499
Asp Asp Lys Gly Cys Pro Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser
635 640 645

ttc aac aac ttc acc gtg agc ttc tgg ctg cgc gtg ccc aag gtg agc 5547
Phe Asn Asn Phe Thr Val Ser Phe Trp Leu Arg Val Pro Lys Val Ser
650 655 660

gcc agc cac ctg gag atc gtc tct gcg gtg gtt ggc att ctg 5589
Ala Ser His Leu Glu Ile Val Ser Ala Val Val Gly Ile Leu
665 670 675

tagaagcttg gtaccgagct cggtccact agtccagtgt ggtggattc tgcagatatc 5649
cagcacatgt gc 5661

<210> 2
<211> 696
<212> PRT
<213> Artificial sequence

<220>
<223> Recombinant expression plasmid derived from pSI

<400> 2

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser
-15 -10 -5

Leu Gly Met Lys His Gln His Gln His Gln His Gln His Gln His Gln
-1 1 5 10

Ala Pro Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys Leu Arg Leu
15 20 25 30

Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His Leu Tyr Gln
35 40 45

Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr Leu Pro Thr
50 55 60

Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val Gln Gly Tyr
65 70 75

Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu Gln Arg Leu
80 85 90

Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr Ala Leu Ala
95 100 105 110

Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro Val Thr Gly
 115 120 125

Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser Leu Thr Glu
 130 135 140

Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln Leu Cys Tyr
 145 150 155

Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn Asn Gln Leu
 160 165 170

Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys His Pro Cys
 175 180 185 190

Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser Ser Glu Asp
 195 200 205

Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys Ala Arg Cys
 210 215 220

Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys Ala Ala Gly
 225 230 235

Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu His Phe Asn
 240 245 250

His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val Gln Tyr Ile
 255 260 265 270

Lys Ala Asn Ser Lys Phe Ile Gly Ile Thr Glu Leu Arg Tyr Thr Phe
 275 280 285

Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu Ser Thr Asp
 290 295 300

Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln Glu Val Thr
 305 310 315

Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys Pro Cys Ala
 320 325 330

Arg Val Cys Tyr Gly Leu Gly Met Glu His Leu Arg Glu Val Arg Ala
 335 340 345 350

Val Thr Ser Ala Asn Ile Gln Glu Phe Ala Gly Cys Lys Lys Ile Phe
355 360 365

Gly Ser Leu Ala Phe Leu Pro Glu Ser Phe Asp Gly Asp Pro Ala Ser
370 375 380

Asn Thr Ala Pro Leu Gln Pro Glu Gln Leu Gln Val Phe Glu Thr Leu
385 390 395

Glu Glu Ile Thr Gly Tyr Leu Tyr Ile Ser Ala Trp Pro Asp Ser Leu
400 405 410

Pro Asp Leu Ser Val Phe Gln Asn Leu Gln Val Ile Arg Gly Arg Ile
415 420 425 430

Leu His Asn Gly Ala Tyr Ser Leu Thr Leu Gln Gly Leu Gly Ile Ser
435 440 445

Trp Leu Gly Leu Arg Ser Leu Arg Glu Leu Gly Ser Gly Leu Ala Leu
450 455 460

Ile His His Asn Thr His Leu Cys Phe Val His Thr Val Pro Trp Asp
465 470 475

Gln Leu Phe Arg Asn Pro His Gln Ala Leu Leu His Thr Ala Asn Arg
480 485 490

Pro Glu Asp Glu Cys Val Gly Glu Gly Leu Ala Cys His Gln Leu Cys
495 500 505 510

Ala Arg Gly His Cys Trp Gly Pro Gly Pro Thr Gln Cys Val Asn Cys
515 520 525

Ser Gln Phe Leu Arg Gly Gln Glu Cys Val Glu Glu Cys Arg Val Leu
530 535 540

Gln Gly Leu Pro Arg Glu Tyr Val Asn Ala Arg His Cys Leu Pro Cys
545 550 555

His Pro Glu Cys Gln Pro Gln Asn Gly Ser Val Thr Cys Phe Gly Pro
560 565 570

Glu Ala Asp Gln Cys Val Ala Cys Ala His Tyr Lys Asp Pro Pro Phe
575 580 585 590

10

Cys Val Ala Arg Cys Pro Ser Gly Val Lys Pro Asp Leu Ser Tyr Met
595 600 605

Pro Ile Trp Lys Phe Pro Asp Glu Glu Gly Ala Cys Gln Pro Cys Pro
610 615 620

Ile Asn Cys Thr His Ser Cys Val Asp Leu Asp Asp Lys Gly Cys Pro
625 630 635

Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser Phe Asn Asn Phe Thr Val
640 645 650

Ser Phe Trp Leu Arg Val Pro Lys Val Ser Ala Ser His Leu Glu Ile
655 660 665 670

Val Ser Ala Val Val Gly Ile Leu
675